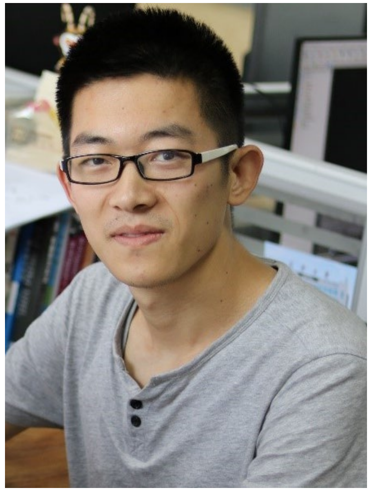
Gaze-Vergence-Controlled See-Through Vision in Augmented Reality

Zhimin Wang[1], Yuxin Zhao[1], Feng Lu[1,2]
1 Beihang University
2 Peng Cheng Laboratory, Shenzhen, China

Project Page: https://zhiminwang.github.io/GVC_See_Through_Vision.html

Authors



Zhimin Wang

Yuxin Zhao

Feng Lu

Outline

- Background
- Related Work
- Our Method
- Experiment
- Limitations and Future Work

Background

- Superman can see objects that are obstructed via superpowers.
- See-through vision: allowing the user to see the occluded objects behind a wall
- Augmented Reality makes this superpower possible.

Superman Clark

Clark’s superpower: see-through vision

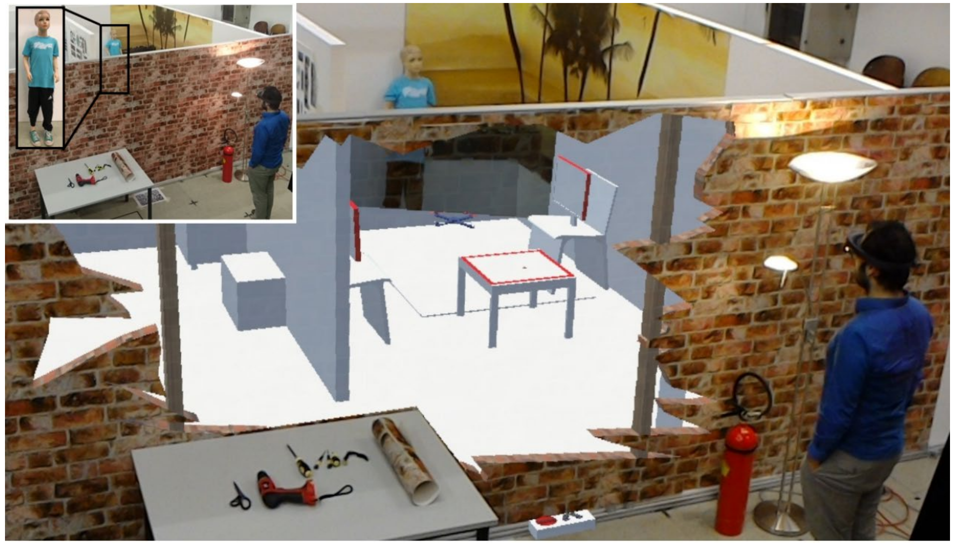
Augmented Reality

Related Work

- Previous literature mainly focused on the overlay effect of hidden areas and occluding layers.
- The way to interact with see-through vision is less studied.

See-through vision with *Edge Overlay* technique

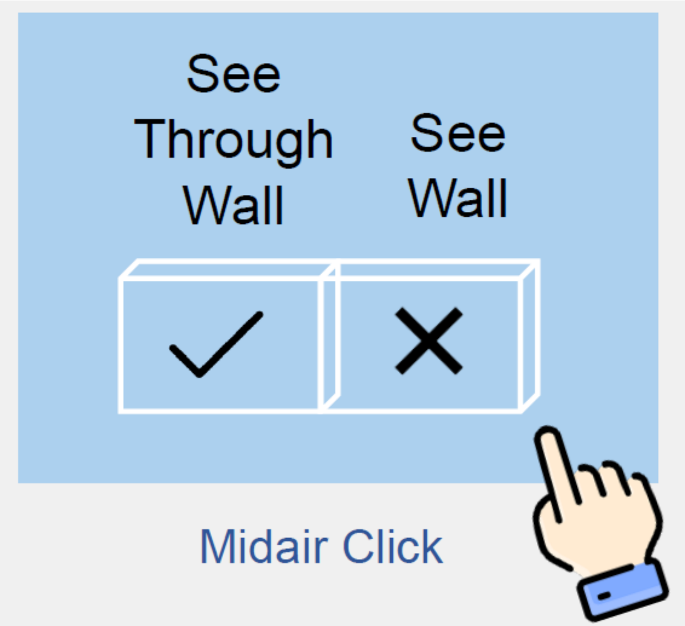
[Avery et al., 2009]

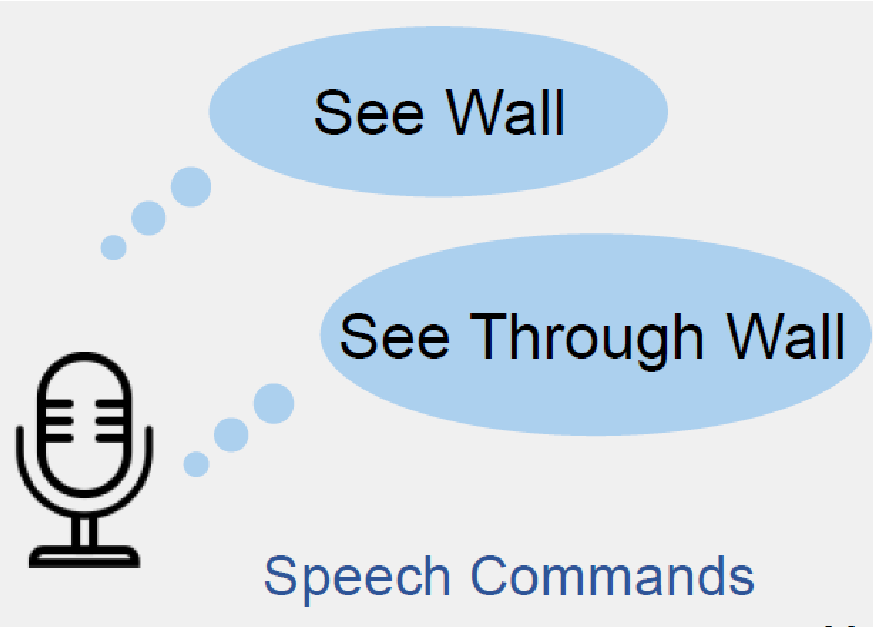


Drone-Augmented Human Vision
[Erat et al. 2018]

Related Work

- Using the common interaction modalities, e.g., midair click and speech, may not be the optimal way to control see-through vision.
- It is not intuitive and requires extra effort to switch the thinking, which will distract the user's attention.





Our Method - Motivation

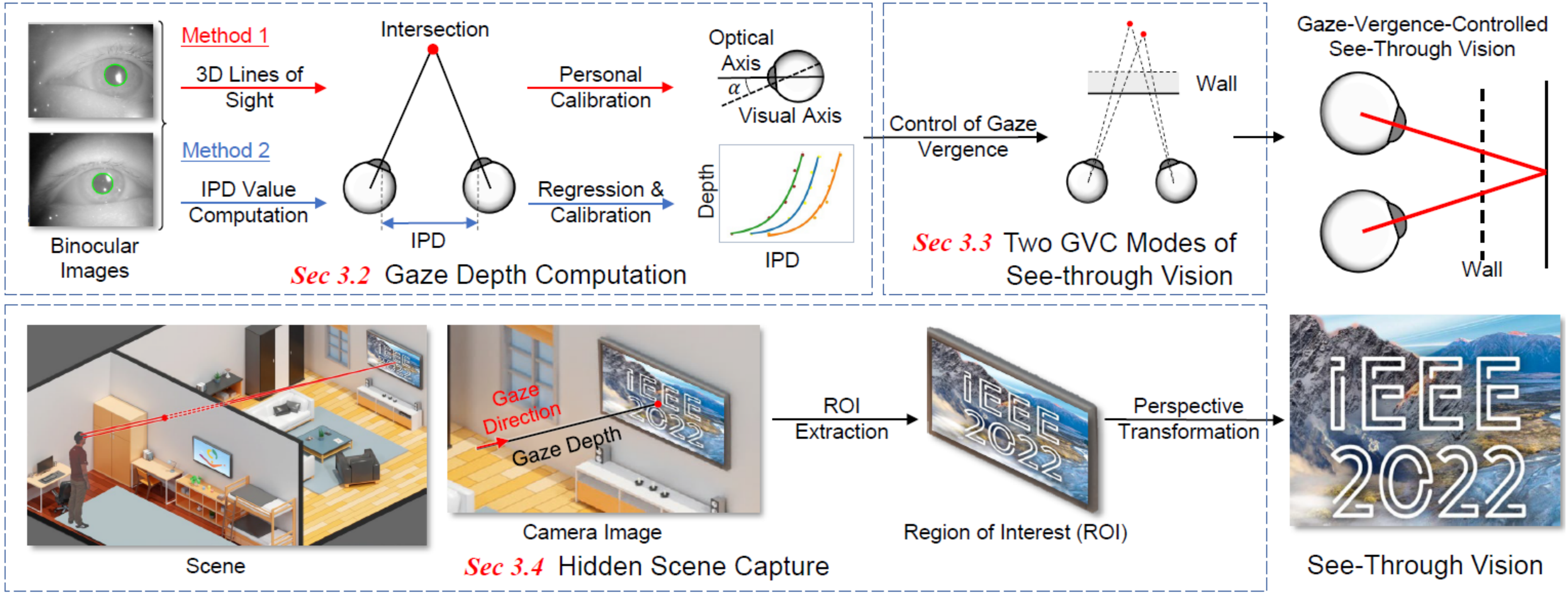
➢Intuitively, when we intend to see through something, we are actually fixating at a new distance, which is physically related to the gaze depth/vergence.

➢We propose a novel gaze-vergence-controlled see-through vision in AR.

- The gaze depth determines whether the see-through vision is triggered.
- The see-through vision's content is determined by the gaze direction + gaze depth.

Our Method - Overview

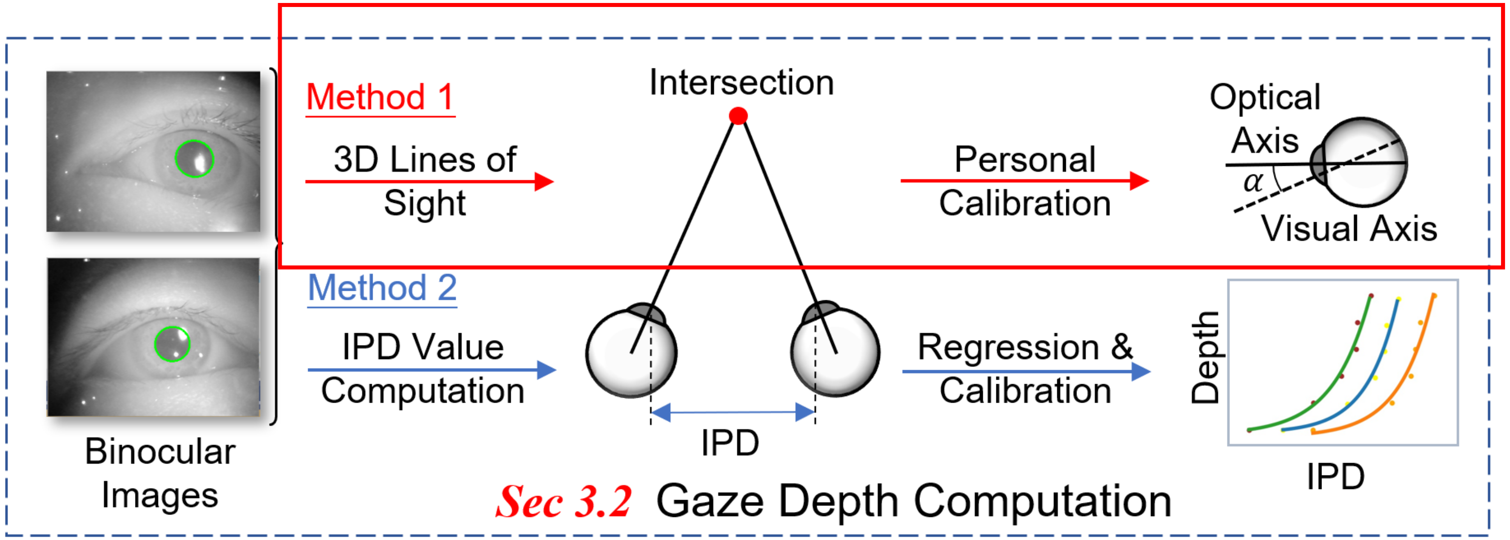
1. We build a gaze tracking module with two infrared cameras and assemble it into the Microsoft HoloLens 2.
2. We design two gaze depth estimation methods, which can be easily adapted to different eye trackers.
3. With our gaze depth estimation algorithm, we propose two control modes of gaze vergence and apply them to see-through vision.

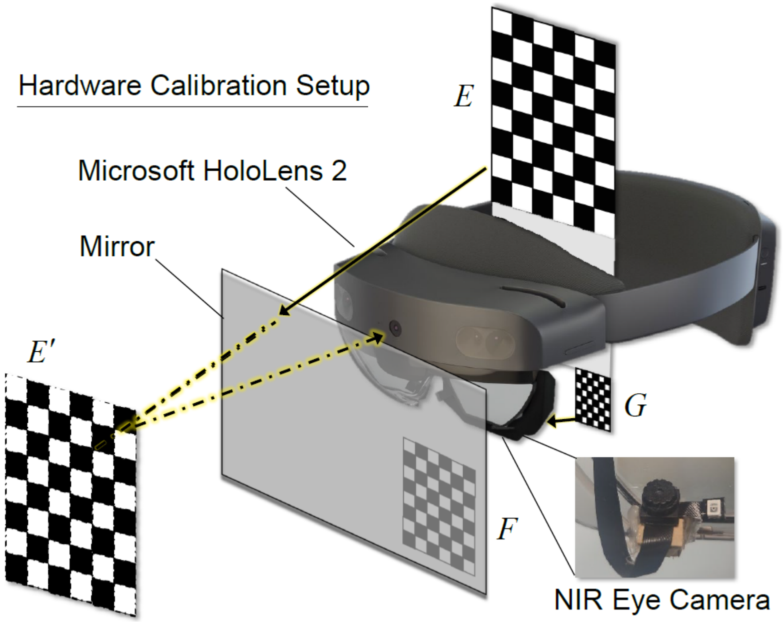


Our Method – Gaze Depth Estimation

We designed two gaze depth computation methods

1. 3D Line-of-sight Intersection (3D LosI)
2. Inter-pupillary Distance (IPD) based Regression





Hardware Calibration Setup

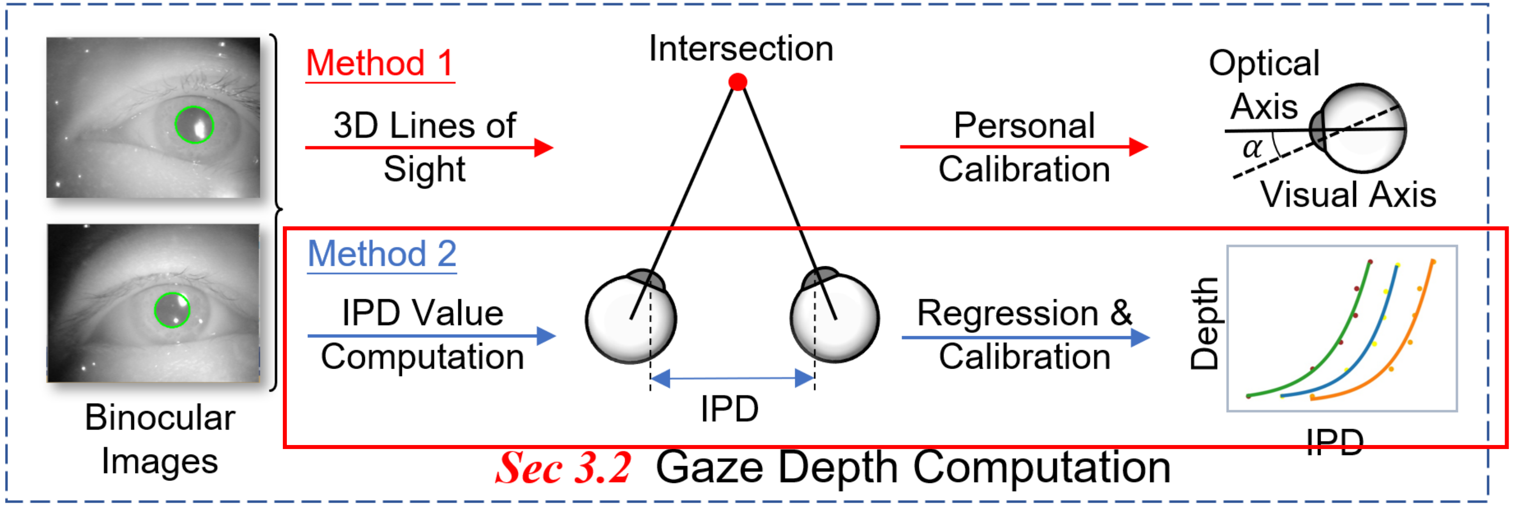
Modified from Pupil Labs' method in two ways:

1. employ the pupil detection method PuReST, which has robust performance to reflections or partial occlusion;
2. calibrate the hardware in advance and model the kappa angle.

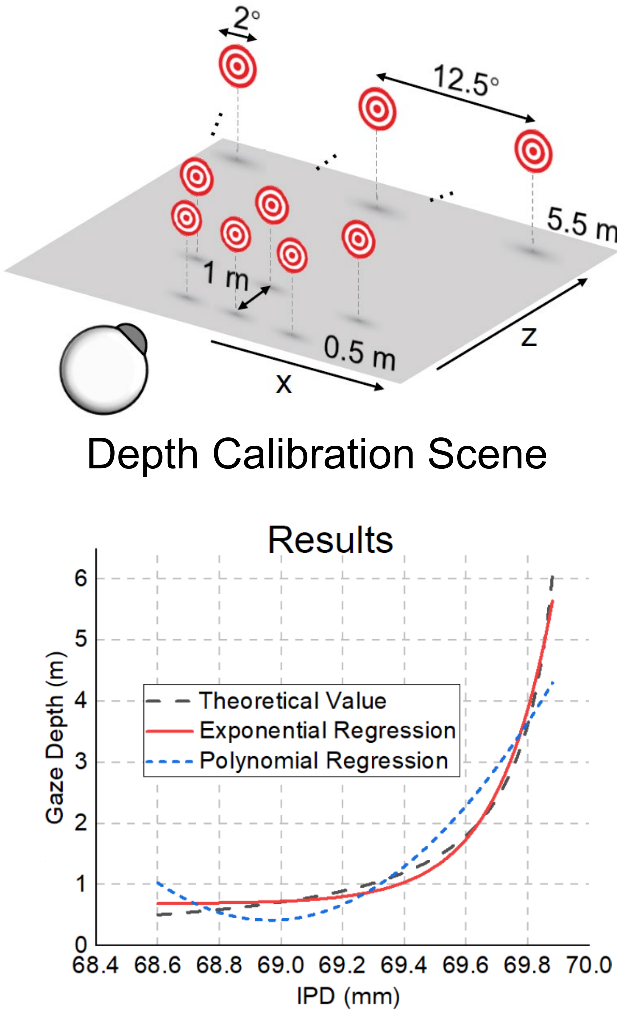
Our Method – Gaze Depth Estimation

We designed two gaze depth computation methods

1. 3D Line-of-sight Intersection (3D LosI)
2. Inter-pupillary Distance (IPD) based Regression



Sec 3.2 Gaze Depth Computation



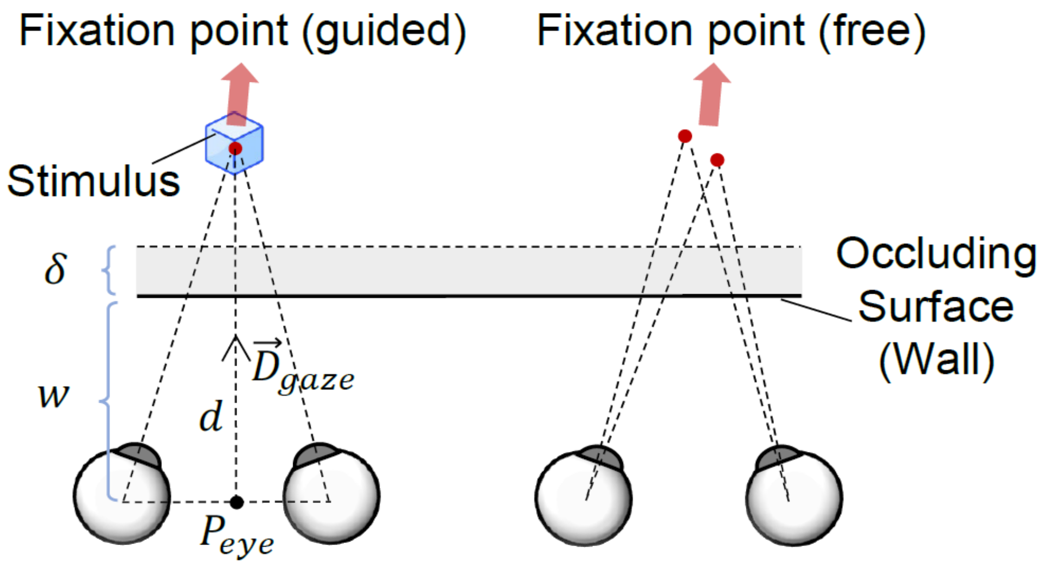
The comparisons among the simulation, the exponential regression and the polynomial regression

we implement two IPD-based methods:

1. utilizes the physical-based IPD in Millimeters (MIPD) to fit gaze depth;
2. uses the image-based IPD in Pixels (PIPD) to regress the depth.

Our Method – Two Control Modes of See-through Vision

- Stimulus-Guided (SG) see-through mode
- Self-Control (SC) see-through mode



SG see-through mode   SC see-through mode

The window position of see-through vision is calculated as:

$$\gamma = \begin{cases} w + j \cdot \delta, & \Phi(d) > w + \delta; \\ -\infty, & \text{otherwise}, \end{cases}$$

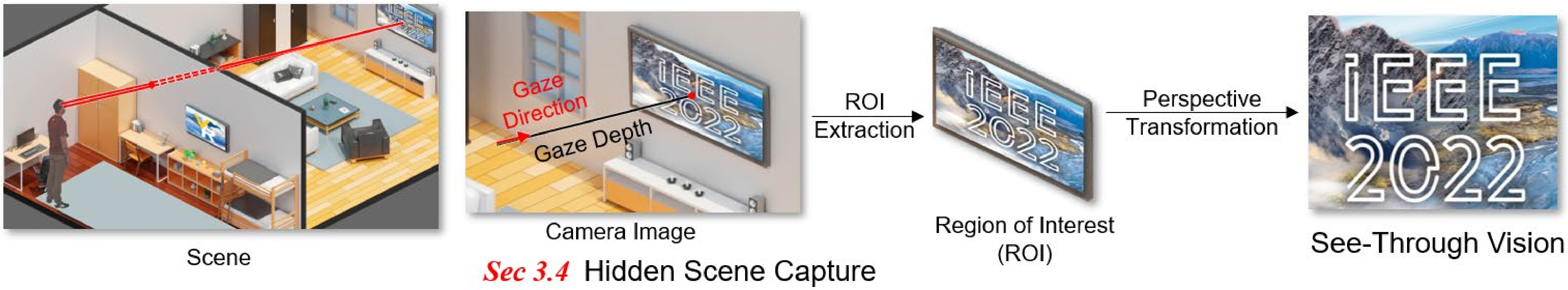
$$P_{window} = P_{eye} + \gamma \cdot \vec{D}_{gaze},$$

γ: the window depth of see-through vision
w: the distance from the user to the wall
δ: the distance threshold
j: a scale factor greater than 1
d: the estimated depth value
$\Phi(\cdot)$: the filter function for data smoothing

P_{window}: the window position of see-through vision
P_{eye}: the center of both eyes
$\vec{D}_{gaze}$: the normal vector of the gaze ray

Our Method - Hidden Scene Capture

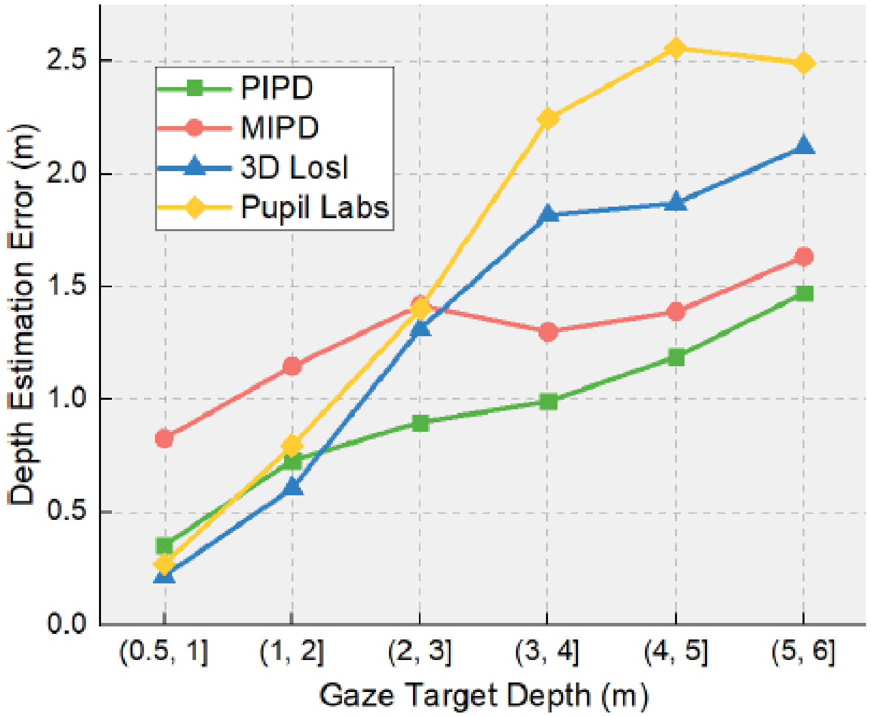
1. Camera Pose Registration: To capture hidden scene, we embed a surveillance camera behind the occluding wall. The camera is first registered to the HoloLens coordinates.
2. ROI Extraction of Hidden Scene: We further compute the ROI in the HoloLens space and map the ROI into the camera space.
3. Perspective Transformation: to make the user's view consistent with physical laws, we apply the perspective transformation method to transform the image of ROI into the user's view in HoloLens.



Experiment – Quantitative evaluation of gaze depth estimation

We evaluated the depth accuracy of our proposed methods, i.e., PIPD, MIPD, 3D LosI, with the Pupil Labs 3D tracker.

- Users: 12 participants
- Distance: (0.5, 6] m



Distance	(0.5, 1]	(1, 2]	(2, 3]	(3, 4]	(4, 5]	(5, 6]
PIPD	0.3±0.3	0.7±0.5	0.9±0.4	1.0±0.5	1.2±0.3	1.5±0.5
MIPD	0.8±1.2	1.1±0.7	1.4±0.8	1.3±0.5	1.4±0.7	1.6±0.6
3D LosI	0.2±0.1	0.6±0.4	1.3±0.9	1.8±0.7	1.9±0.4	2.1±0.4
Pupil Labs	0.3±0.2	0.8±0.5	1.4±0.7	2.2±0.6	2.6±0.5	2.5±0.4

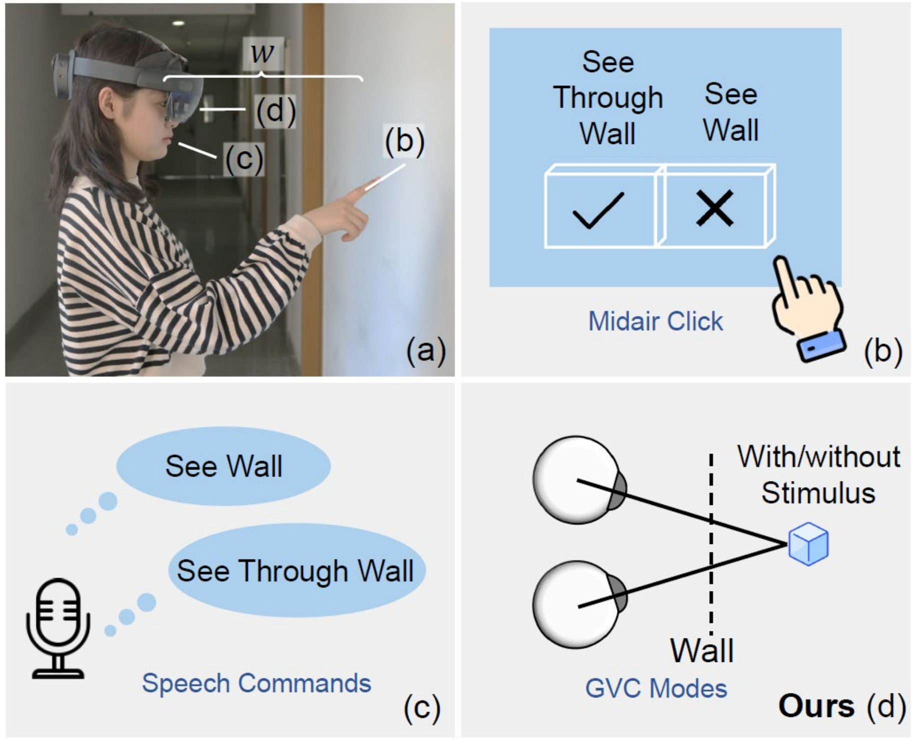
Results:
1) 3D LosI achieves the best performance in the range of (0.5, 2] m;
2) The PIPD outperforms the other methods at the (2, 6] m

Discussion:
- combine the 3D LosI and PIPD for gaze vergence control.
- Use the gaze vergence to perform daily indoor interaction within the middle distance, i.e., (0.5, 3] m.

Experiment – Comparisons of modalities for see-through vision control

We compare the Gaze-Vergence-Controlled (GVC) techniques with two common modalities.

- Users: 20 participants
- Distance: 1, 2, 3 m
- Four techniques: Stimulus-Guided Gaze(**SGGaze**), Self-Control Gaze (**SCGaze**), midair click technique (**Click**) and speech-based technique (**Speech**)
- Performance Measures
  - Completion Time
  - The number of successes
  - The number of mistakes
- Subjective Measures
  - NASA's Task Load Index
  - User preference

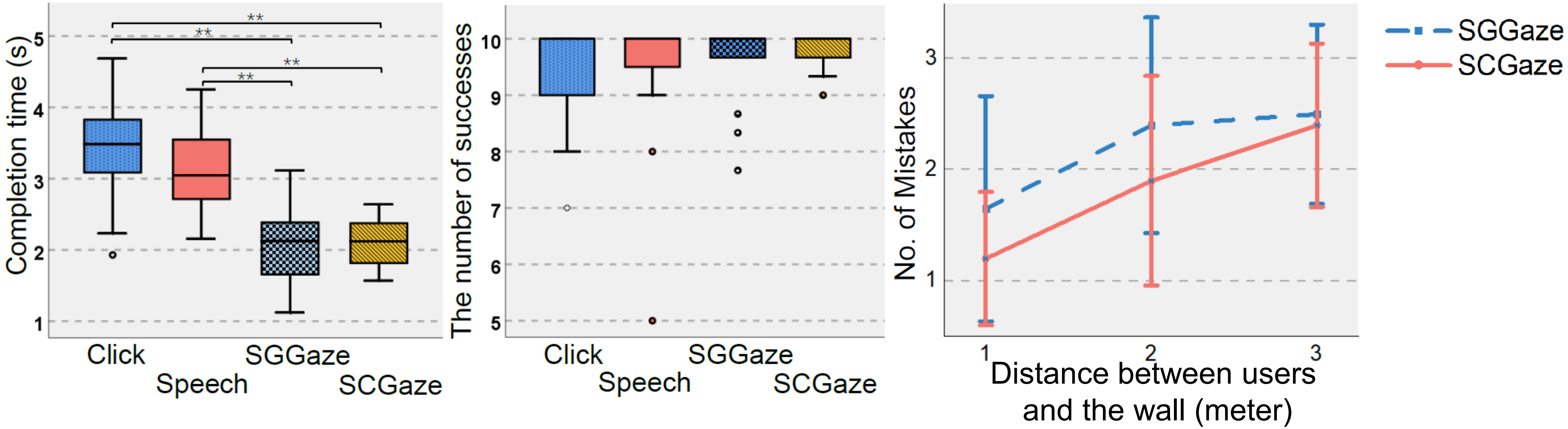


Four interaction modalities for see-through vision control

Experiment – Comparisons of modalities for see-through vision control

Performance Measures

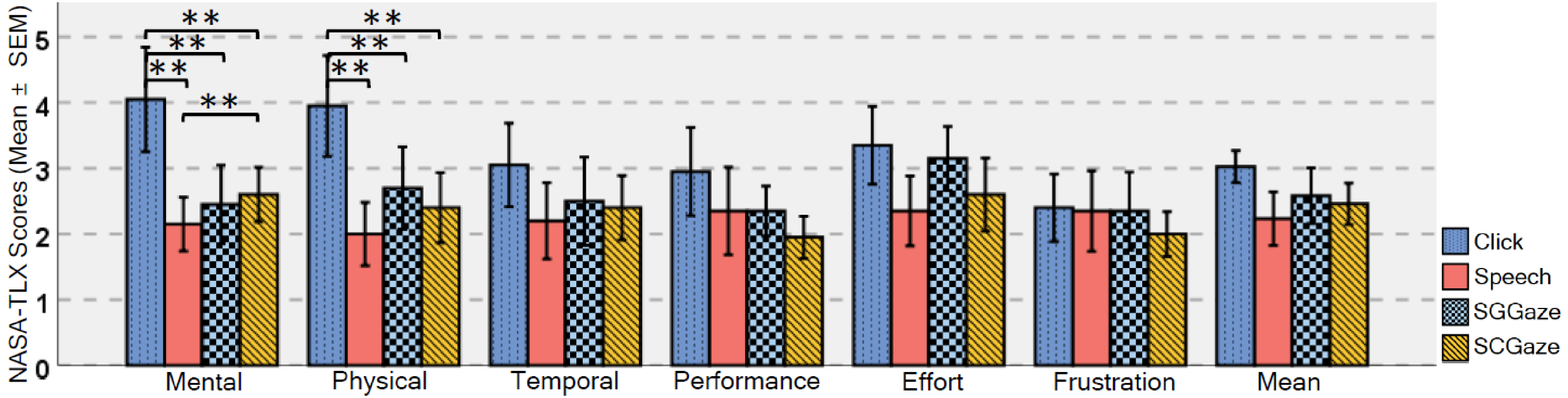
- SGGaze and SCGaze were significantly faster than the two common modalities ($p < 0.001$, 0.001).
- Users can almost finish the correct operations at the assigned time.
- The number of mistakes increased with increasing distances for GVC techniques.



Experiment – Comparisons of modalities for see-through vision control

Subjective Measures

- The Click achieved the highest mental/physical demand than all the other techniques.
- The Speech had lower mental demand than the SCGaze.
- There is no significant difference in terms of other task loads.



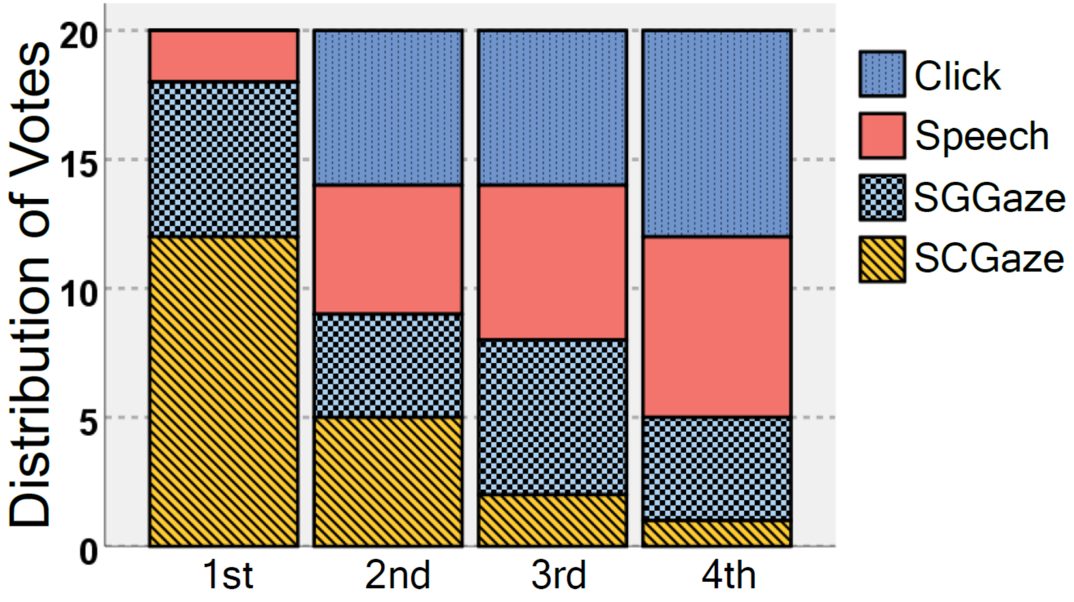
Bar charts of scores on the NASA-TLX questionnaire for comparing four modalities

Experiment – Comparisons of modalities for see-through vision control

Subjective Measures

The SCGaze is the most preferred by the users.

- P3: *" I feel arm fatigue after Click. "*
- P4: *" The speech command needs to speak aloud to trigger the switch, which is not convenient in a quiet space. "*
- P16*: " It is amazing. I have been looking in the same direction, but the change of vergence can convey a signal of seeing through the wall, which is a novel experience for me. "*



The user preference ranking of four interaction modalities

Limitations and Future Work

- It is difficult to discriminate the vergence difference when the distance exceeds 3 m. Therefore, for long-distance interaction (>3 m), we can use the modality independent of the distance, e.g., speech-based technique.

- In the future, we will design and implement a shared see-through vision between multiple users controlled by gaze vergence.

Thank you!